DOCUMENT RESUME

ED 202 688 SE 034 921

AUTHOR Robinson, James T.
TITLE Preparing and Using Data Tapes and Machine Readable User's Guides: A New Resource for Science Education Research.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE Apr 81
GRANT NSF-SED-79-19312
NOTE 28p.; Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (54th, Grossinger's in the Catskills, Ellenville, NY, April 5-8, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Data Analysis; *Data Bases; *Data Collection; *Data Processing; *Documentation; Higher Education; Information Storage; *Research Methodology; *Science Education; Secondary Education; Secondary School Science
IDENTIFIERS National Science Foundation; *Science Education Research

ABSTRACT

The preparation of a data tape and users guide is reported. The tape contains about 700 cases with nearly two thousand variables collected over a three-year period as part of a curriculum evaluation study. The student group is the eleven- to fourteen-year-old population that participated in the field test of the BSCS Human Sciences Program. (Author/JN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED202688

CERE CENTER FOR EDUCATIONAL RESEARCH AND EVALUATION
P.O. BOX 930 BOULDER, COLORADO 80306 PHONE (303) 666-6558

833 W. South Boulder Road
Louisville, Colorado 80027

ADVISORY COMMITTEE
J. Myron Atkin
Stanford University
Stanford
William P. Callahan
University of Northern Iowa
Cedar Falls
Jack A. Easley, Jr.
University of Illinois
Urbana-Champaign
William Kastrinos
Educational Testing Service
Princeton
Oliver P. Kolstoe
University of
Northern Colorado
Greeley
Jon D. Miller
Northern Illinois University
DeKalb
James P. Shaver
Utah State University
Logan
Lorrie Shepard
University of Colorado
Boulder
Robert S. Siegler
Carnegie-Mellon University
Pittsburgh
Barbara Tymitz
Indiana University
Bloomington
Herbert J. Walberg
University of Illinois
at Chicago Circle
Chicago
Wayne W. Welch
University of Minnesota
Minneapolis
Jon Wiles
University of Montana
Missoula

Preparing and Using Data Tapes and Machine Readable User's Guides: A New Resource for Science Education Research

James T. Robinson

Center for Educational Research and Evaluation, BSCS

Louisville, Colorado

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

JAMES T. ROBINSON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A Paper Presented at The

National Association for Research in Science Teaching

April, 1981

SE 034 921

The research reported here was supported by Grant No. SED 79-19312 from the Research in Science Education (RISE) program of the National Science Foundation. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author and do not necessarily reflect the views of the NSF.

Introduction

The purpose of this paper is to introduce novices to the preparation and uses of machine readable data files (MRDF). A fully documented machine readable data file consists of two basic units: a computer tape or other storage medium carrying the machine readable data, and documentation, a tape or hard copy document that serves to communicate information about the data in order to permit users, uninvolved in the data collection process to gain access to the data file and to use all or part of the data for their own research or to compare with their research.

The central elements of developing MRDF's will be illustrated with examples from this investigator's current work in which data collected from 1973 to 1977 are being prepared as a MDRF using the Statistical Package for the Social Sciences (SPSS) as the computer software for the task.

The literature for preparation of machine readable data files is limited and largely fugitive. Most of the work done has been in the social sciences, library archive preparation, or in work with data management systems in business and industry. The work reported here, and it is not as yet completed, relies most heavily on Robbin (1974-75) and Roistacher (1979). The final product will include the major elements from these sources, adapted to the special needs of science education and to the SPSS software.

Values and Uses of MRDF's

Most research in science education is conducted by investigators gathering their own data, albeit on a small group of subjects, and analyzing their own results. More recently (Glass, 1978) has advocated the use of meta-analysis to combine data from many published studies as a useful research tool. He demonstrated the usefulness of meta-analysis in research on psychotherapy (Glass and Smith, 1976) and the relationship of class size to achievement. Other researchers have utilized meta-analysis procedures advantageously since Glass's procedures were published.

Published summary statistics have limitations that also become limitations in meta-analysis procedures. Natural and social sciences researchers have long accumulated data, stored these in computer tapes, and made them available to the research community. Data tapes, such as those made available by government agencies such as the Census Bureau have been used extensively as data sources for research. The costs of securing valid data are increasing, and therefore, more attention needs to be given to data documentation, data sharing, and the multiple use of such a resource. Yet, this kind of resource is not readily available for science education research.

Coordinated research programs developed and conducted by an individual investigator and her or his graduate students or by a group of investigators and students can be enhanced and facilitated by the preparation and execution of plans for developing and managing data files. The limitations of isolated empirical studies are obvious to most investigators, but turnover of personnel and the lack of detailed documentation of work completed is an added deterrent to interrelating studies for the development of cumulative knowledge in science education.

Secondary analyses, made feasible by well documented data, have made invaluable contributions to knowledge in the social and natural science disciplines. They can be equally useful to the complex field of science education.

Terminology

The terminology related to machine readable data files is not yet completely standardized, but the terminology explained and adopted here is in conformity with Federal Information Processing Standards for documentation.

Machine readable refers to files of numeric data in rectangular form and the documentation and formatting of those data.

Data file is the file of numeric data recorded in rectangular form on magnetic tape or disc, or in columns on cards.

User's guide is a comprehensive manual that describes the data file's identity, organization, contents, physical characteristics, and relation to computer hardware and software (Roistacher, 1979). For convenience, the user's guide can be subdivided into three parts: source documentation, codebook documentation, and technical documentation. (Robbin, 1974-75).

Source documentation is a description of the study and its procedures.

Codebook documentation describes the logical structure of data items, their location in the collection of information, and the values ascribed to each data item (Robbin, 1974-75). Roistacher (1979) refers to the codebook as the "data dictionary."

Technical documentation provides the information needed by users to access the file at their computing facility.

These terms will be further elaborated as the development of a MRDF is described.

Data File

The data file can be produced by using a variety of computer programs. The Statistical Package for the Social Sciences (SPSS) offers several advantages over other systems in this writer's judgment. The advantages of SPSS are the extensive labeling of variables and the value of variables, the availability of "comments" that can be included in the file, the extensive range of statistical analyses that can be computed with the package, and the provision for file manipulation, data transformation, and file management. SPSS is an integrated system of computer programs for the analysis of the kinds of data of importance in science education. The package enables the science educator to perform analyses through the use of natural-language control statements. Space does not permit an extensive discussion of the SPSS. Interested readers are referred to Nie, et al (1975) and Hull and Nie (1979) for complete information on SPSS. The ensuing discussion is based on the use of SPSS as a computer package for developing machine readable data files.

The data file is prepared and stored as an SPSS system file. This procedure results in a file in which all the variables can be named and labeled, the values of each variable can be named and labeled, and comments can be interspersed as needed.

The structure of an SPSS system file is based on cases, each with associated variables and values. All of the variables and values for each case are associated with identifiers for the case. Cases can be added to or deleted from the file. Variables can be added or deleted and new variables computed for each case and saved as part of the file. Cases can be sorted and variables can be rearranged within the file. (See Appendix A for an example.)

A critical phase in the preparation of data files is the process of checking and verifying the accuracy of the data before entering them

in the SPSS system file. Once the data are in the SPSS system file, the user can run several kinds of accuracy checks. For example, by using the FREQUENCIES program for all ID characters and for each variable, codings that are outside designated values can be easily spotted. Table 1 shows the output of such a run in the sex coding in a file. One keypunch error, a "3"

Table 1. Frequencies for the two values of the sex code in File EXPYR3.

SEX 1=MALE 2=FEMALE

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
MALE	1	349	48.1	48.1	48.1
FEMALE	2	376	51.8	51.9	100.0
OUT OF RANGE		1	.1	MISSING	
	TOTAL	726	100.0	100.0	

was exposed. It showed up as an "out of range" value in Table 1. This error could be corrected laboriously by hand or, if it was found that the correct code should have been a "1," the "3" could be permanently recorded as a "1," using the appropriate IF statement.

To determine whether the variables and values for each case are accurate, the SAMPLE card can be used to specify a percentage of the cases for random sampling from the file. A random sample can be run using the LIST CASES facility, specifying all the variables for the cases. The output from this run can be checked to ascertain the completeness and accuracy of the file.

User's Guide

The user's guide should include all of the information needed by any researcher to utilize the data with integrity and without the necessity for discussions with the file producer. For convenience of discussion and of

preparation and production, the user's guide can be described in three sections: source, codebook, and technical documentation.

Source Documentation

Source documentation describes the study or situation during which the data were gathered. It includes the theoretical framework, study design, all procedures and instruments used, how editing and data cleanup were accomplished, and how missing data, ambiguities and marking errors were resolved. Any decisions made that affect the data, its completeness and quality, need to be provided to the user.

If interview data or coded data are included, instructions, procedures, protocols, reliability, and other pertinent data are to be reported in the user's guide. Definitions of variables, how variables were generated, any data transformations and how they were produced are important kinds of information, and also important to the potential user.

Full disclosure of procedures used to assure respondent privacy and confidentiality, if applicable, should be provided. It may be that some data were aggregated to preserve confidentiality. Users should be cautioned about disaggregations that could violate confidentiality of respondents (Roistacher, 1979).

Finally, if statistical output was generated, it is important to report how it was produced. If results of analyses using data in the file have been published, a bibliography of publications should be included in the user's guide.

Source document material, since it is mostly prose, is generated using a text editor such as SCRIBE, TEXT EDITOR, and/or PROSE. The plan initially formulated for the current project was to compose this material using a text editor via time-sharing. Word processing equipment with telecommunications capability now seems preferable to time-sharing for this task because of the reduced cost due to loss on-line time.

Codebook Documentation

According to Robbin (1974-75) the ideal codebook should "stand alone" by describing every variable contained in the data file. Codebook documentation includes the full text of questions, test items, interview questions and complete variable and value names for each data item. Ranges, means, and standard deviations should be provided for each variable as well as frequencies for each value of each variable.

The SPSS subroutines, CONDESCRIPTIVE or FREQUENCIES, can be used to calculate the statistics for each variable. FREQUENCIES is most useful because it provides the number and percent of respondents selecting each value of each variable and descriptive statistics for the variable. A reference dictionary listing the variables alphabetically and by their sequence in the file is an option that can be output using the FREQUENCIES facility. Appendix B illustrates the three types of outputs availabe from FREQUENCIES described here.

Technical Documentation

This documentation provides potential users with the names of each file and the tape record specifications for each file: number of tracks on the tape (seven or nine track), density (BPI), parity, record blocking, maximum block size, whether the tape is labeled, and the character code (ASCII or EBCDIC). In most instances, the potential user should contact her or his computing facility before ordering a data tape to determine how the tape files should be written for her or his computer facility as the tape output to the recipient can be written to those specifications. Appendix C provides an example of an order form that a data file producer would provide to a potential user so that a machine readable data tape and user's guide could be properly produced.

Additional Assets of An SPSS System File

Use of an SPSS system file requires that the user have this package available at her or his computer facility. Without this capability the data file and a printed user's guide can make the data available. However, the documentation, variable labels, variable names and value labels will not be available as part of a machine readable data record.

With an SPSS capability, the data file can be output as card images in the character code required by the user. This is accomplished by using the WRITE CASES procedure in SPSS. Labels and comments are output by using the LIST FILEINFO or WRITE FILEINFO procedures. WRITE FILEINFO can produce either punched cards or a tape with the card images for most of the data documentation in the SPSS system file. The format statement from the WRITE CASES statement should be supplied to the user. This can be easily adapted to the INPUT FORMAT statement for reading the data from the cards or tape.

LIST FILEINFO would be used alone when sending data documentation to a non-SPSS user. LIST FILEINFO would be used together with WRITE FILEINFO to produce a hard copy document of the contents of the cards or tape produced by the WRITE FILEINFO procedure.

If cases need to be deleted or selected from SPSS system files, two procedures are available to use in conjunction with the SAVE FILE function of SPSS. For deleting cases, a REJECT IF statement is available on the 8.0 and later version of SPSS. This allows the case by case deletion from the files to be saved. Using the SELECT IF together with the NE (not equal) operator produces logic errors that result in no cases being deleted. REJECT IF together with the EQ (equal) operator will properly delete any cases identified.

The SELECT IF statement, used together with the EQ (equal) operator can select any subgroup of cases for any analysis desired. If the number of cases

is large, the REJECT IF statement can be used together with the EQ operator to delete a small number of cases. While a complete SPSS system file cannot be transferred directly from one computer to another, the procedures listed above will enable the user to reconstitute the SPSS system file on her or his own hardware.

Files with up to 500 variables can be processed by the standard SPSS program. Files of up to 1,000 variables can be processed by the large version of SPSS. Files of up to 5,000 variables can be produced by using the ARCHIVE procedures of SPSS. In addition to the large number of statistical procedures available in SPSS, a newly added REPORT capability (Hull and Nie, 1979) enables users to format tables for publications directly from the SPSS system file. Variables can be called by names from a system file without the need for specifying format information.

Producing Machine Readable Data Files

The use of the Statistical Package for the Social Sciences (SPSS) has been briefly described here as a system of computer programs that has advantages for producing MRDFs. This system plus text editing programs and the necessary hardware are all that are needed.

A word processor that can be used as an interactive terminal offers a much more efficient and economical means for developing the user's guide. The prose sections can be stored in floppy discs of the word processor. Any statistical tables, for example, those for the codebook, could be generated using the SPSS and then disposed from the computer file to the word processor files. Editing and rearranging, and the merger of prose and tables can be easily accomplished on the word processor, building the final user's guide, which can then be disposed to and stored in the computer system files.

Several options will be available for providing potential users with the data file and user's guide: both furnished on computer tape; data file on

tape, user's guide printed out by either computer or word processor; data file and/or user's guide transmitted by telephone to potential user's word processor.

Summary

The purpose of this paper is to recommend the preparation and use of machine readable data files as a valuable tool for research in science education. The quality and usability of such files is dependent on the collection, verification, and documentation of such files. Once carefully prepared, these data can be used for secondary analysis, for original research in asking new questions of the data, and as comparison data for numerous studies.

Documentation is of critical importance. It serves four important functions for data by providing:

1. a conceptual framework
2. historical reference
3. general instructions for communication between the producer and user
4. reports of successful output (Robbin, 1974-75).

I wish to acknowledge the comments and suggestions made by my colleague, Richard R. Tolman, on a previous draft of this paper. Many of his ideas were incorporated into this final paper.

References Cited

Glass, Gene V. and Smith, M. L. Meta-analysis of psychotherapy outcome studies. Paper presented at the Annual Meeting of the Society for Psychotherapy Research, San Diego, CA., June, 1976.

Glass, Gene V. Integrating findings: the meta-analysis of research. In L. S. Shulman (ed.). Review of Research in Education, (Vol. 5). Itasca, IL: Peacock, 1978.

Hull, C. Hadlai and Nie, Norman H. (eds.) SPSS Update, New Procedures and Facilities for Releases 7 and 8. New York: McGraw Hill, 1979.

Nie, Norman H., Hull, C. Hadlai, Jenkins, Jean G., Steinbrenner, Karin, and Bent, Dale H. SPSS, Statistical Package for the Social Sciences. 2nd Edition. New York: McGraw Hill, 1975.

Robbin, Alice. Managing information access through documentation of the data base: characterizing social science data base text documentation as a minimal information management system. Sigsoc Bullecin 6, 2-3, 56-68, Fall-Winter, 1974-75.

Roistacher, Richard C. A style manual for machine readable data files and their documentation, draft 5 of 16:13:09, 08-20-79. Bureau of Social Science Research, Washington, D. C.: January, 1979. Unpublished manuscript.

Addendum

Roistacher, Richard C., with Dodd, S. M. Noble, B. B., and Robbin, A. A style manual for machine-readable data files and their documentaion. U. S. Department of Justice, Bureau of Justice Statistics, Superintendent of Documents, U. S. Government Printing Office, Washington, D. C., June, 1980.

APPENDIX "A"

CONTENTS OF CASE NUMBER 29

SEQNUM	29.	SUBFILE	HSPALL	CASWGT	1.0000	PERSON	29.	STATUS1	0
SEX	2.	SCH74	1.	TEAC74	1.	SNR74	43.	STATUS2	0
SCH75	1.	TEAC75	2.	SNR75	43.	SCH76	1.	TEAC76	2.
SNR76	43.	AGE	146.	VAR008	7.	VAR009	3.	VAR010	6.
VAR011	1.	VAR012	3.	VAR013	3.	VAR014	1.	VAR015	1.
VAR016	3.	VAR017	1.	VAR018	1.	VAR019	2.	VAR020	1.
VAR021	2.	VAR022	3.	VAR023	3.	VAR024	2.	VAR025	4.
VAR026	3.	VAR027	6.	VAR028	3.	VAR029	3.	VAR030	6.
VAR031	6.	VAR032	6.	VAR033	6.	VAR034	6.	VAR035	6.
VAR036	6.	VAR037	6.	VAR038	6.	VAR039	6.	VAR040	3.
VAR041	3.	VAR042	6.	VAR043	3.	VAR044	1.	VAR045	2.
VAR046	2.	VAR047	1.	VAR048	1.	VAR049	5.	VAR050	1.
VAR051	1.	VAR052	2.	VAR053	2.	VAR054	3.	VAR055	1.
VAR056	4.	VAR057	5.	VAR058	4.	VAR059	5.	VAR060	5.
VAR061	1.	VAR062	1.	VAR063	1.	VAR064	1.	VAR065	1.
VAR066	5.	VAR067	4.	VAR068	1.	VAR069	1.	VAR070	1.
VAR071	5.	VAR072	5.	VAR073	1.	VAR074	3.	VAR075	3.
VAR076	3.	VAR077	3.	VAR078	5.	VAR079	5.	VAR080	2.
VAR081	5.	VAR082	4.	VAR083	3.	VAR084	4.	VAR085	3.
VAR086	1.	VAR087	0	VAR088	58.	VAR089	39.	VAR090	85.
VAR091	47.	VAR092	16.	VAR400	0	VAR401	0	VAR402	1.
VAR403	1.	VAR404	0	VAR405	0	VAR406	1.	VAR407	0
VAR408	0	VAR409	3.	VAR410	1.	VAR411	1.	REP001	0
REP002	0	REP003	0	REP004	0	REP005	0	REP006	0
REP007	0	REP008	0	REP009	0	REP010	0	REP011	0
REP012	0	REP013	0	REP014	0	REP015	0	REP016	0
REP017	0	REP018	0	REP019	0	REP020	0	REP021	0
REP022	0	REP023	0	REP024	0	REP025	0	REP026	0
REP027	0	REP028	0	REP029	0	REP030	0	REP031	0
REP032	0	REP033	0	REP034	0	REP035	0	REP036	0
REP037	0	REP038	0	REP039	0	REP101	0	REP102	0
REP103	0	REP104	0	REP105	0	REP106	0	REP107	0
REP108	0	REP109	0	REP110	0	REP111	0	REP112	0
REP113	0	REP114	0	REP115	0	REP116	0	REP117	0
REP118	0	REP119	0	REP120	0	REP121	0	REP122	0
REP123	0	REP124	0	REP125	0	REP126	0	REP127	0
REP128	0	REP129	0	REP130	0	REP131	0	REP132	0
REP133	0	REP134	0	REP135	0	REP136	0	REP137	0
REP138	0	REP139	0	REP140	0	REP141	0	REP201	0
REP202	0	REP203	0	REP204	0	REP205	0	REP206	0
REP207	0	REP208	0	REP209	0	REP210	0	REP211	0
REP212	0	REP213	0	REP214	0	REP215	0	REP216	0
REP217	0	REP218	0	REP219	0	REP220	0	REP301	0
REP302	0	REP303	0	REP304	0	REP305	0	REP306	0
REP307	0	REP308	0	REP309	0	REP310	0	REP311	0
REP312	0	REP313	0	REP314	0	REP315	0	REP316	0
REP317	0	REP318	0	REP319	0	REP320	0	REP321	0

CONTENTS OF CASE NUMBER 30

11442/0013

SEQNUM	30.	SUBFILE	HSPALL	CASWGT	1.0000	PERSON	30.	STATUS1	0
SEX	2.	SCH74	1.	TEAC74	1.	SNR74	44.	STATUS2	0
SCH75	1.	TEAC75	3.	SNR75	44.	SCH76	0	TEAC76	0
SNR76	0	AGE	0	VAR008	0	VAR009	0	VAR010	0
VAR011	0	VAR012	0	VAR013	0	VAR014	0	VAR015	0
VAR016	0	VAR017	0	VAR018	0	VAR019	0	VAR020	0
VAR021	0	VAR022	0	VAR023	0	VAR024	0	VAR025	0
VAR026	0	VAR027	0	VAR028	0	VAR029	0	VAR030	0
VAR031	0	VAR032	0	VAR033	0	VAR034	0	VAR035	0
VAR036	0	VAR037	0	VAR038	0	VAR039	0	VAR040	0
VAR041	0	VAR042	0	VAR043	0	VAR044	0	VAR045	0
VAR046	0	VAR047	0	VAR048	0	VAR049	0	VAR050	0
VAR051	0	VAR052	0	VAR053	0	VAR054	0	VAR055	0
VAR056	0	VAR057	0	VAR058	0	VAR059	0	VAR060	0
VAR061	0	VAR062	0	VAR063	0	VAR064	0	VAR065	0
VAR066	0	VAR067	0	VAR068	0	VAR069	0	VAR070	0
VAR071	0	VAR072	0	VAR073	0	VAR074	0	VAR075	0
VAR076	0	VAR077	0	VAR078	0	VAR079	0	VAR080	0
VAR081	0	VAR082	0	VAR083	0	VAR084	0	VAR085	0
VAR086	0	VAR087	0	VAR088	0	VAR089	0	VAR090	0
VAR091	0	VAR092	0	VAR400	0	VAR401	0	VAR402	0
VAR403	0	VAR404	0	VAR405	0	VAR406	0	VAR407	0
VAR408	0	VAR409	0	VAR410	0	VAR411	0	REP001	0
REP002	0	REP003	0	REP004	0	REP005	0	REP006	0

APPENDIX "B"

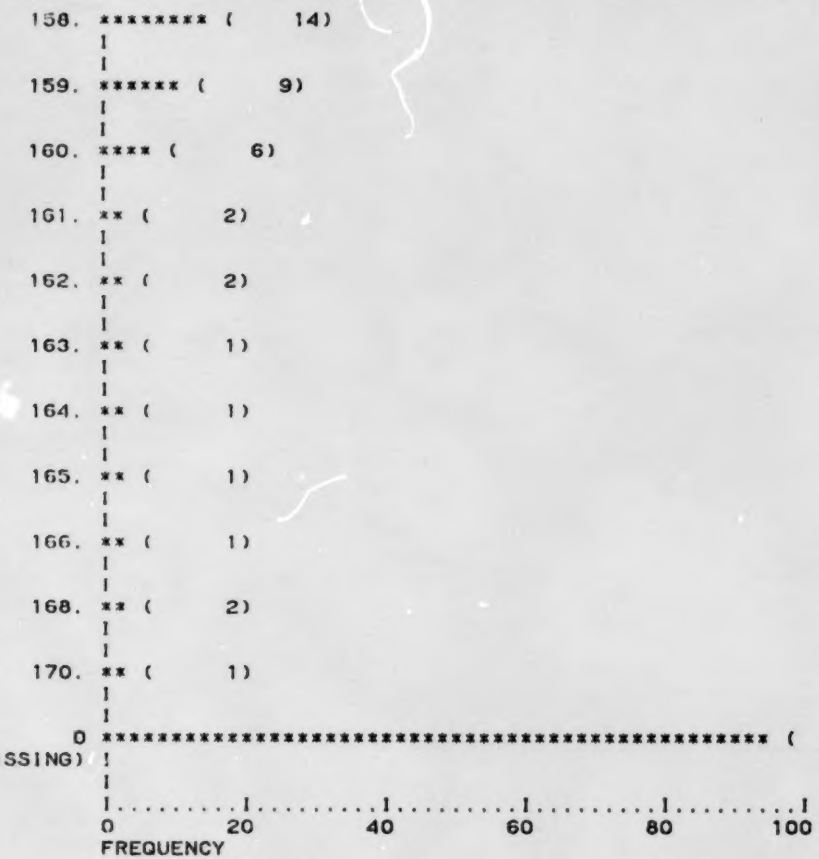
AGE AGE IN MONTHS AS OF MAY 1974

CATEGORY LABEL	CODE	ABSOLUTE FREQ	RELATIVE FREQ (PCT)	ADJUSTED FREQ (PCT)	CUM FREQ (PCT)
9/case 72081, changed to 150	114.	1	.1	.1	.1
	122.	1	.1	.1	.3
	124.	1	.1	.1	.4
	126.	4	.5	.6	1.0
	127.	2	.2	.3	1.3
	128.	1	.1	.1	1.4
	129.	1	.1	.1	1.6
	130.	3	.4	.4	2.0
	133.	1	.1	.1	2.1
	134.	2	.2	.3	2.4
	136.	5	.6	.7	3.1
	137.	25	3.1	3.5	6.6
	138.	22	2.8	3.1	9.8
	139.	38	4.7	5.4	15.1
	140.	56	7.0	7.9	23.1
	141.	52	6.5	7.4	30.4
	142.	51	6.4	7.2	37.6
	143.	44	5.5	6.2	43.8
	144.	37	4.6	5.2	49.1
	145.	48	6.0	6.8	55.9
	146.	43	5.4	6.1	62.0
	147.	42	5.3	5.9	67.9
	148.	47	5.9	6.6	74.5
	149.	37	4.6	5.2	79.8

FILE EXPYR1 (CREATION DATE = 81/02/20.) ADDED AGES,DELETED AND RENAMED VARIABLES 1973-74

150.	26	3.3	3.7	83.5
151.	23	2.9	3.3	86.7
152.	12	1.5	1.7	88.4
153.	4	.5	.6	89.0
154.	9	1.1	1.3	90.2
155.	7	.9	1.0	91.2
156.	8	1.0	1.1	92.4
157.	14	1.7	2.0	94.3
158.	14	1.7	2.0	96.3
159.	9	1.1	1.3	97.6
160.	6	.7	.8	98.4
161.	2	.2	.3	98.7
162.	2	.2	.3	99.0
163.	1	.1	.1	99.2
164.	1	.1	.1	99.3
165.	1	.1	.1	99.4
166.	1	.1	.1	99.6
168.	2	.2	.3	99.9
170.	1	.1	.1	100.0
0	93	11.6	MISSING	
TOTAL	800	100.0	100.0	

FILE EXPYR1 (CREATION DATE = 81/02/20.) ADDED AGES,DELETED AND RENAMED VARIABLES 1973-74



MEAN	145.181	STD ERR	.248	MEDIAN	144.635
MODE	140.000	STD DEV	6.594	VARIANCE	43.486
KURTOSIS	1.711	SKEWNESS	.275	RANGE	56.000
MINIMUM	114.000	MAXIMUM	170.000	SUM	102643.000
C.V. PCT	4.542	.95 C.I.	144.694	TO	145.668

VALID CASES 707 MISSING CASES 93

FILE EXPYR3 (CREATION DATE = 81/02/26.) 1975-76 HSP DATA FILE

POSITIONAL INDEX

VARIABLE	PAGE	VARIABLE	PAGE	VARIABLE	PAGE	VARIABLE	PAGE
SEX	2	CHG202	24	FEL008	45	FEL029	66
AGE	3	CHG203	25	FEL009	46	FEL030	67
CHG101	5	CHG204	26	FEL010	47	FEL031	68
CHG102	6	CHG205	27	FEL011	48	FEL032	69
CHG103	7	CHG206	28	FEL012	49	FEL033	70
CHG104	8	CHG207	29	FEL013	50	FEL034	71
CHG105	9	CHG208	30	FEL014	51	FEL035	72
CHG106	10	CHG209	31	FEL015	52	FEL036	73
CHG107	11	CHG210	32	FEL016	53	FEL037	74
CHG108	12	CHG211	33	FEL017	54	FEL038	75
CHG109	13	CHG212	34	FEL018	55	FEL039	76
CHG110	14	CHG213	35	FEL019	56	FEL040	77
CHG111	15	CHG214	36	FEL020	57	FEL041	78
CHG112	16	CHG215	37	FEL021	58	FEL042	79
CHG113	17	FEL001	38	FEL022	59	FEL043	80
CHG114	18	FEL002	39	FEL023	60	FEL044	81
CHG115	19	FEL003	40	FEL024	61	FEL045	82
CHG116	20	FEL004	41	FEL025	62	FEL046	83
CHG117	21	FEL005	42	FEL026	63		
CHG118	22	FEL006	43	FEL027	64		
CHG201	23	FEL007	44	FEL028	65		

FILE EXPYR3 (CREATION DATE = 81/02/26.) 1975-76 HSP DATA FILE

ALPHABETIC INDEX

VARIABLE	PAGE	VARIABLE	PAGE	VARIABLE	PAGE	VARIABLE	PAGE
AGE	3	CHG203	25	FEL009	46	FEL030	67
CHG101	5	CHG204	26	FEL010	47	FEL031	68
CHG102	6	CHG205	27	FEL011	48	FEL032	69
CHG103	7	CHG206	28	FEL012	49	FEL033	70
CHG104	8	CHG207	29	FEL013	50	FEL034	71
CHG105	9	CHG208	30	FEL014	51	FEL035	72
CHG106	10	CHG209	31	FEL015	52	FEL036	73
CHG107	11	CHG210	32	FEL016	53	FEL037	74
CHG108	12	CHG211	33	FEL017	54	FEL038	75
CHG109	13	CHG212	34	FEL018	55	FEL039	76
CHG110	14	CHG213	35	FEL019	56	FEL040	77
CHG111	15	CHG214	36	FEL020	57	FEL041	78
CHG112	16	CHG215	37	FEL021	58	FEL042	79
CHG113	17	FEL001	38	FEL022	59	FEL043	80
CHG114	18	FEL002	39	FEL023	60	FEL044	81
CHG115	19	FEL003	40	FEL024	61	FEL045	82
CHG116	20	FEL004	41	FEL025	62	FEL046	83
CHG117	21	FEL005	42	FEL026	63	SEX	2
CHG118	22	FEL006	43	FEL027	64		
CHG201	23	FEL007	44	FEL028	65		
CHG202	24	FEL008	45	FEL029	66		

EXAMPLE 14: DATA ORDER FORM , from Roistacher (1980)

Name: _____

Address: _____

_____

_____

Tape Recording Specifications

Seven-track Tape

Density (BPI)	200	556	800
Parity	Even	Odd	
Record blocking	Blocked	Unblocked	
Maximum block size	_____		
IBM standard labels	Labeled	Not labeled	
1-6 character label	_____		
Character code	BCD		

Nine-track Tape

Density (BPI)	800	1600	6250
Parity	Odd		
Record blocking	Blocked	Unblocked	
Maximum block size	_____		
IBM standard labels	Labeled	Not labeled	
1-6 character label	_____		
Character code	EBCDIC		

(Over)

Example 14: Data Order Form